

(1) GENERAL INFORMATION:

(ii) TITLE OF INVENTION: Bacterial Superantigen Vaccines

(iv) CORRESPONDENCE ADDRESS:

(B) STREET: US Army MPMC -504 Scott Street
MCMR-JA (Elizabeth Arwine-Patent
Atty)

(D) STATE: MARYLAND

(E) COUNTRY: USA

(F) ZIP: 21702-5012

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: Apple Macintosh

(C) OPERATING SYSTEM: Macintosh 7.5

(D) SOFTWARE: Microsoft Word 6.0

(A) APPLICATION NUMBER:

(B) FILING DATE:

(C) CLASSIFICATION:

(A) APPLICATION NUMBER: 08/882,431

(B) FILING DATE: June 25, 1997

(A) NAME: Sana A. Pratt

(B) REGISTRATION NUMBER: 39,441

(C) REFERENCE/DOCKET NUMBER:

(A) TELEPHONE: (301) 619-2065

(B) TELEFAX: (301) 619-7714

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 830

(B) TYPE: Nucleic Acid

(C) STRANDEDNESS: Unknown

(D) TOPOLOGY: Unknown

(D) TOPOLOGY: Unknown

(ii) Molecule type: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGAAAAAAAA CAGCATTTAC ATTACTTTTA TTCATTGCCC	40
TAACGTTGAC AACAAGTCCA CTTGTAAATG GTAGCGAGAA	80
AAGCGAAGAA ATAAATGAAA AAGATTTGCG AAAAAAGTCT	120
GAATTGCAGG GAACAGCTTT AGGCAATCTT AAACAAATCT	160
ATTATTACAA TGAAAAAGCT AAAACTGAAA ATAAAGAGAG	200
TCACGATCAA TTTCGACAGC ATACTATATT GTTTAAAGGC	240
TTTTTTACAG ATCATTCGTG GTATAACGAT TTATTAGTAC	280
GTTTTGATTC AAAGGATATT GTTGATAAAT ATAAAGGGAA	320
AAAAGTAGAC TTGTATGGTG CTTATGCTGG TTATCAATGT	360
GCGGGTGGTA CACCAAACAA AACAGCTTGT ATGTATGGTG	400
GTGTAACGTT ACATGATAAT AATCGATTGA CCGAAGAGAA	440
AAAAGTGCCG ATCAATTTAT GGCTAGACGG TAAACAAAAT	480
ACAGTACCTT TGGAAACGGT TAAAACGAAT AAGAAAAATG	520
TAAGTGTTCG GGAGTTGGAT CTTCAAGCAA GACGTTATTT	560
ACAGGAAAAA TATAATTTAT ATAAGTCTGA TGTTTTTGAT	600
GGGAAGGTTT AGAGGGGATT AATCGTGTTT CATACTTCTA	640
CAGAACCTTC GGTTAATTAC GATTTATTTG GTGCTCAAGG	680
ACAGTATTCA AATACACTAT TAAGAATATA TAGAGATAAT	720
AAAACGATTA ACTCTGAAAA CATGCATATT GATATATATT	760
TATATACAAG TTAAACATGG TAGTTTTGAC CAACGTAATG	800
TTCAGATTAT TATGAACCGA GAATAATCTA	830

(3) INFORMATION FOR SEQUENCE ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 257

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

(ii) Molecule type: Peptide

Met Lys Lys Thr Ala Phe Thr Leu Leu Leu
1 5 10

Leu Val Asn Gly Ser Glu Lys Ser Glu Glu
25 30

Glu Leu Gln Gly Thr Ala Leu Gly Asn Leu
45 50

Lys Thr Glu Asn Lys Glu Ser His Asp Gln
65 70

Phe Phe Thr Asp His Ser Trp Tyr Asn Asp
85 90

Val Asp Lys Tyr Lys Gly Lys Lys Val Asp
105 110

Ala Gly Gly Thr Pro Asn Lys Thr Ala Cys
125 130

Asn Arg Leu Thr Glu Glu Lys Lys Val Pro
145 150

Ile Asn Leu Trp Leu Asp Gly Lys Gln Asn
155 160

Thr Val Pro Leu Glu Thr Val Lys Thr Asn
 165 170
 Lys Lys Asn Val Thr Val Gln Glu Leu Asp
 175 180
 Leu Gln Ala Arg Arg Tyr Leu Gln Glu Lys
 185 190
 Tyr Asn Leu Tyr Asn Ser Asp Val Phe Asp
 195 200
 Gly Lys Val Gln Arg Gly Leu Ile Val Phe
 205 210
 His Thr Ser Thr Glu Pro Ser Val Asn Tyr
 215 220
 Asp Leu Phe Gly Ala Gln Gly Gln Tyr Ser
 225 230
 Asn Thr Leu leu Arg Ile Tyr Arg Asp Asn
 235 240
 Lys Thr Ile Asn Ser Glu Asn Met His Ile
 245 250
 Asp Ile Tyr Leu Tyr Thr Ser
 255

(4) INFORMATION FOR SEQUENCE ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 757
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Unknown
- (D) TOPOLOGY: Unknown

(ii) Molecule type: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGAGAAAAG CGAAGAAATA AATGAAAAAG ATTTGCGAAA 40
 AAAGTCTGAA TTGCAGGGAA CAGCTTTAGG CAATCTTAAA 80
 CAAATCTATT ATTACAATGA AAAAGCTAAA ACTGAAAATA 120
 AAGAGAGTCA CGATCAATTT CGACAGCATA CTATATTGTT 160
 TAAAGGCTTT TTTACAGATC ATTCGTGGTA TAACGATTTA 200

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

(5) INFORMATION FOR SEQUENCE ID NO:4:

(A) LENGTH: 233

(B) TYPE: Amino Acid

(C) STRANDEDNESS: Unknown

(D) TOPOLOGY: Unknown

(ii) Molecule type: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Glu Lys Ser Glu Glu Ile Asn Glu Lys
5 10

Asp Leu Arg Lys Lys Ser Glu Leu Gln Gly
15 20

Thr Ala Leu Gly Asn Leu Lys Gln Ile Tyr
25 30

Tyr Tyr Asn Glu Lys Ala Lys Thr Glu Asn
35 40

Lys Glu Ser His Asp Gln Phe Arg Gln His
45 50

Thr	Ile	Leu	Phe	Lys	Gly	Phe	Phe	Thr	Asp	
				55					60	
His	Ser	Trp	Tyr	Asn	Asp	Leu	Leu	Val	Arg	
				65					70	
Phe	Asp	Ser	Lys	Asp	Ile	Val	Asp	Lys	Tyr	
				75					80	
Lys	Gly	Lys	Lys	Val	Asp	Leu	Tyr	Gly	Ala	
				85					90	
Tyr	Ala	Gly	Tyr	Gln	Cys	Ala	Gly	Gly	Thr	
				95					100	
Pro	Asn	Lys	Thr	Ala	Cys	Met	Tyr	Gly	Gly	
				105					110	
Val	Thr	Leu	His	Asp	Asn	Asn	Arg	Leu	Thr	
				115					120	
Glu	Glu	Lys	Lys	Val	Pro	Ile	Asn	Leu	Trp	
				125					130	
Leu	Asp	Gly	Lys	Gln	Asn	Thr	Val	Pro	Leu	
				135					140	
Glu	Thr	Val	Lys	Thr	Asn	Lys	Lys	Asn	Val	
				145					150	
Thr	Val	Gln	Glu	Leu	Asp	Leu	Gln	Ala	Arg	
				155					160	
Arg	Tyr	Leu	Gln	Glu	Lys	Tyr	Asn	Leu	Tyr	
				165					170	
Asn	Ser	Asp	Val	Phe	Asp	Gly	Lys	Val	Gln	
				175					180	
Arg	Gly	Leu	Ile	Val	Phe	His	Thr	Ser	Thr	
				185					190	
Glu	Pro	Ser	Val	Asn	Tyr	Asp	Leu	Phe	Gly	
				195					200	
Ala	Gln	Gly	Gln	Tyr	Ser	Asn	Thr	Leu	Leu	
				205					210	
Arg	Ile	Tyr	Arg	Asp	Asn	Lys	Thr	Ile	Asn	
				215					220	

1002784 12504

Tyr Thr Ser

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1712
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Unknown
(D) TOPOLOGY: Unknown

(ii) Molecule type: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GAAC TAGGTA	GAAAAATAAT	TATGAGAAAA	CACTATGTTG	40
TTAAAGATGT	TTTCGTATAT	AAGTTTAGGT	GATGTATAGT	80
TACTTAATTT	TAAAAGCATA	ACTTAATTAA	TATAAATAAC	120
ATGAGATTAT	TAAATATAAT	TAAGTTTCTT	TTAATGTTTT	160
TTTAATTGAA	TATTTAAGAT	TATAACATAT	ATTTAAAGTG	200
TATCTAGATA	CTTTTTGGGA	ATGTTGGATA	AAGGAGATAA	240
AAAATGTATA	AGAGATTATT	TATTTCACAT	GTAATTTTGA	280
TATTCGCACT	GATATTAGTT	ATTTCTACAC	CCAACGTTTT	320
AGCAGAGAGT	CAACCAGATC	CTAAACCAGA	TGAGTTGCAC	360
AAATCGAGTA	AATTCACTGG	TTTGATGGAA	GATATGAAAG	400
TTTTGTATGA	TGATAATCAT	GTATCAGCAA	TAAACGT TAA	440
ATCTATAGAT	CAATTTCTAT	ACTTTGACTT	AATATATTCT	480
ATTAAGGACA	CTAAGTTAGG	GGATTATGAT	AATGTTTCGAG	520
TCGAATTTAA	AAACAAAGAT	TTAGCTGATA	AATACAAAGA	560
TAAATACGTA	GATGTGTTTG	GAGCTAATTA	TTATTATCAA	600
TGTTATTTTT	CTAAAAAAC	GAATGATATT	AATTCGCATC	640
AAACTGACAA	ACGAAAAACT	TGTATGTATG	GTGGTGTAAC	680

TGAGCATAAT	GGAAACCAAT	TAGATAAAATA	TAGAAGTATT	720
ACTGTTCTGGG	TATTTGAAGA	TGGTAAAAAT	TTATTATCTT	760
TTGACGTACA	AACTAATAAG	AAAAAGGTGA	CTGCTCAAGA	800
ATTAGATTAC	CTAACTCGTC	ACTATTTGGT	GAAAAATAAA	840
AAACTCTATG	AATTTAACAA	CTCGCCTTAT	GAAACGGGAT	880
ATATTAAATT	TATAGAAAAT	GAGAATAGCT	TTTGGTATGA	920
CATGATGCCT	GCACCAGGAG	ATAAATTTGC	CCAATCTAAA	960
TATTTAATGA	TGTACAATGA	CAATAAAATG	GTTGATTCTA	1000
AAGATGTGAA	GATTGAAGTT	TATCTTACGA	CAAAGAAAAA	1040
GTGAAATTAT	ATTTTAGAAA	AGTAAATATG	AAGAGTTAGT	1080
AATTAAGGCA	GGCACTTATA	GAGTACCTGC	CTTTTCTAAT	1120
ATTATTTAGT	TATAGTTATT	TTTGTTATAT	CTCTCTGATT	1160
TAGCATTAAAC	CCCTTGTTGC	CATTATAGTT	TTCACCAACT	1200
TTAGCTGAAA	TTGGGGGATC	ATTTTTATCT	TTACTATGGA	1240
TAGTTACTGT	GTCGCCGTTT	TTAACGATTT	GTTTCTCTTT	1280
TAATTTGTCA	GTTAATTTT	TCCATGCATC	ATTTGCGTCA	1320
AACCTATTTT	CATTTGGATT	TATCTTTGAC	AAATCAATTC	1360
TTTTAACACT	ATCGGTATTA	ATCGGCTTGT	TATTAAAATT	1400
ACTAAGTTCA	TCTAAATCAG	CTGTACCCGT	AATACTACTT	1440
TCGCCACCAT	TATTTAAATT	GTACGTAACA	CCAACGTCT	1480
CATTTGCTGT	TTTATCGATA	ATATTTGCTT	CTTTCAAAGC	1520
ATCTCTTACA	TTTTTCCATA	AGTCTCTATC	TGTTATTTCA	1560
GAAGCCTTTG	CAACGTTATT	AATACCATTA	TAATTTGAAG	1600
AAGAATGAAA	ACCTGAACCT	ACTGTTGTTA	AAACTAAAGC	1640
ACTTGCTATC	AATGTTCTTG	TTAATAGTTT	TTTATTCATT	1680
TTATTTTCTC	CTATAACTTA	TTTGCAATCG	AT	1712

(i) SEQUENCE CHARACTERISTICS:

(B) TYPE: Amino Acid

(D) TOPOLOGY: Unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Ser His Gln Thr Asp Lys Arg Lys Thr Cys
135 140

Met	Tyr	Gly	Gly	Val 145	Thr	Glu	His	Asn	Gly 150
Asn	Gln	Leu	Asp	Lys 155	Tyr	Arg	Ser	Ile	Thr 160
Val	Arg	Val	Phe	Glu 165	Asp	Gly	Lys	Asn	Leu 170
Leu	Ser	Phe	Asp	Val 175	Gln	Thr	Asn	Lys	Lys 180
Lys	Val	Thr	Ala	Gln 185	Glu	Leu	Asp	Tyr	Leu 190
Thr	Arg	His	Tyr	Leu 195	Val	Lys	Asn	Lys	Lys 200
Leu	Tyr	Glu	Phe	Asn 205	Asn	Ser	Pro	Tyr	Glu 210
Thr	Gly	Tyr	Ile	Lys 215	Phe	Ile	Glu	Asn	Glu 220
Asn	Ser	Phe	Trp	Tyr 225	Asp	Met	Met	Pro	Ala 230
Pro	Gly	Asp	Lys	Phe 235	Ala	Gln	Ser	Lys	Tyr 240
Leu	Met	Met	Tyr	Asn 245	Asp	Asn	Lys	Met	Val 250
Asp	Ser	Lys	Asp	Val 255	Lys	Ile	Glu	Val	Tyr 260
Leu	Thr	Thr	Lys	Lys 265	Lys				

(8) INFORMATION FOR SEQUENCE ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1712
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Unknown
(D) TOPOLOGY: Unknown

(ii) Molecule type: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAACTAGGTA GAAAATAAT TATGAGAAAA CACTATGTTG

[illegible]

ATTATTTAGT TATAGTTATT TTTGTTATAT CTCTCTGATT 1160
 TAGCATTAAAC CCCTTGTTGC CATTATAGTT TTCACCAACT 1200
 TTAGCTGAAA TTGGGGGATC ATTTTATCT TTAATATGGA 1240
 TAGTTACTGT GTCGCCGTTT TTAACGATTT GTTCTCTCTT 1280
 TAATTTGTCA GTTAATTTTT TCCATGCATC ATTTGCGTCA 1320
 AACCTATTTT CATTGGAATT TATTCTTGAC AAATCAATTC 1360
 TTTTAACACT ATCGGTATTA ATCGGCTTGT TATTAAAATT 1400
 ACTAAGTTCA TCTAAATCAG CTGTACCCGT AATACTACTT 1440
 TCGCCACCAT TATTTAAATT GTACGTAACA CCAACTGTCT 1480
 CATTTGCTGT TTTATCGATA ATATTTGCTT CTTTCAAAGC 1520
 ATCTCTTACA TTTTCCATA AGTCTCTATC TGTTATTTCA 1560
 GAAGCCTTTG CAACGTTATT AATACCATTA TAATTTGAAG 1600
 AAGAATGAAA ACCTGAACCT ACTGTTGTTA AAATAAAGC 1640
 ACTTGCTATC AATGTTCTTG TTAATAGTTT TTTATTCATT 1680
 TTATTTTCTC CTATAACTTA TTTGCAATCG AT 1712

(9) INFORMATION FOR SEQUENCE ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 266
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Unknown
- (D) TOPOLOGY: Unknown

(ii) Molecule type: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met	Tyr	Lys	Arg	Leu	Phe	Ile	Ser	His	Val
				5					10
Ile	Leu	Ile	Phe	Ala	Leu	Ile	Leu	Val	Ile
				15					20
Ser	Thr	Pro	Asn	Val	Leu	Ala	Glu	Ser	Gln
				25					30

Pro	Asp	Pro	Lys	Pro	Asp	Glu	Leu	His	Lys	
				35					40	
Ser	Ser	Lys	Phe	Thr	Gly	Leu	Met	Glu	Asn	
				45					50	
Met	Lys	Val	Leu	Tyr	Asp	Asp	Asn	His	Val	
				55					60	
Ser	Ala	Ile	Asn	Val	Lys	Ser	Ile	Asp	Gln	
				65					70	
Phe	Arg	Tyr	Phe	Asp	Leu	Ile	Tyr	Ser	Ile	
				75					80	
Lys	Asp	Thr	Lys	Leu	Gly	Asn	Tyr	Asp	Asn	
				85					90	
Val	Arg	Val	Glu	Phe	Lys	Asn	Lys	Asp	Leu	
				95					100	
Ala	Asp	Lys	Tyr	Lys	Asp	Lys	Tyr	Val	Asp	
				105					110	
Val	Phe	Gly	Ala	Asn	Ala	Tyr	Tyr	Gln	Cys	
				115					120	
Ala	Phe	Ser	Lys	Lys	Thr	Asn	Asp	Ile	Asn	
				125					130	
Ser	His	Gln	Thr	Asp	Lys	Arg	Lys	Thr	Cys	
				135					140	
Met	Tyr	Gly	Gly	Val	Thr	Glu	His	Asn	Gly	
				145					150	
Asn	Gln	Leu	Asp	Lys	Tyr	Arg	Ser	Ile	Thr	
				155					160	
Val	Arg	Val	Phe	Glu	Asp	Gly	Lys	Asn	Leu	
				165					170	
Leu	Ser	Phe	Asp	Val	Gln	Tyr	Asn	Lys	Lys	
				175					180	
Lys	Val	Thr	Ala	Gln	Glu	Leu	Asp	Tyr	Leu	
				185					190	
Thr	Arg	His	Tyr	Leu	Val	Lys	Asn	Lys	Lys	
				195					200	
Leu	Tyr	Glu	Phe	Asn	Asn	Ser	Pro	Tyr	Glu	
				205					210	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

[illegible][illegible][illegible]

- [illegible]

α_1^{II}	α_1^{III}	α_1^{IV}	α_1^{V}	α_1^{VI}	α_1^{VII}	α_1^{VIII}	α_1^{IX}	α_1^{X}	α_1^{XI}	α_1^{XII}
β_1^{I}	β_1^{II}	β_1^{III}	β_1^{IV}	β_1^{V}	β_1^{VI}	β_1^{VII}	β_1^{VIII}	β_1^{IX}	β_1^{X}	β_1^{XI}
γ_1^{I}	γ_1^{II}	γ_1^{III}	γ_1^{IV}	γ_1^{V}	γ_1^{VI}	γ_1^{VII}	γ_1^{VIII}	γ_1^{IX}	γ_1^{X}	γ_1^{XI}
δ_1^{I}	δ_1^{II}	δ_1^{III}	δ_1^{IV}	δ_1^{V}	δ_1^{VI}	δ_1^{VII}	δ_1^{VIII}	δ_1^{IX}	δ_1^{X}	δ_1^{XI}
ϵ_1^{I}	ϵ_1^{II}	ϵ_1^{III}	ϵ_1^{IV}	ϵ_1^{V}	ϵ_1^{VI}	ϵ_1^{VII}	ϵ_1^{VIII}	ϵ_1^{IX}	ϵ_1^{X}	ϵ_1^{XI}
ζ_1^{I}	ζ_1^{II}	ζ_1^{III}	ζ_1^{IV}	ζ_1^{V}	ζ_1^{VI}	ζ_1^{VII}	ζ_1^{VIII}	ζ_1^{IX}	ζ_1^{X}	ζ_1^{XI}
η_1^{I}	η_1^{II}	η_1^{III}	η_1^{IV}	η_1^{V}	η_1^{VI}	η_1^{VII}	η_1^{VIII}	η_1^{IX}	η_1^{X}	η_1^{XI}
θ_1^{I}	θ_1^{II}	θ_1^{III}	θ_1^{IV}	θ_1^{V}	θ_1^{VI}	θ_1^{VII}	θ_1^{VIII}	θ_1^{IX}	θ_1^{X}	θ_1^{XI}
ι_1^{I}	ι_1^{II}	ι_1^{III}	ι_1^{IV}	ι_1^{V}	ι_1^{VI}	ι_1^{VII}	ι_1^{VIII}	ι_1^{IX}	ι_1^{X}	ι_1^{XI}
κ_1^{I}	κ_1^{II}	κ_1^{III}	κ_1^{IV}	κ_1^{V}	κ_1^{VI}	κ_1^{VII}	κ_1^{VIII}	κ_1^{IX}	κ_1^{X}	κ_1^{XI}
λ_1^{I}	λ_1^{II}	λ_1^{III}	λ_1^{IV}	λ_1^{V}	λ_1^{VI}	λ_1^{VII}	λ_1^{VIII}	λ_1^{IX}	λ_1^{X}	λ_1^{XI}
μ_1^{I}	μ_1^{II}	μ_1^{III}	μ_1^{IV}	μ_1^{V}	μ_1^{VI}	μ_1^{VII}	μ_1^{VIII}	μ_1^{IX}	μ_1^{X}	μ_1^{XI}
ν_1^{I}	ν_1^{II}	ν_1^{III}	ν_1^{IV}	ν_1^{V}	ν_1^{VI}	ν_1^{VII}	ν_1^{VIII}	ν_1^{IX}	ν_1^{X}	ν_1^{XI}
ξ_1^{I}	ξ_1^{II}	ξ_1^{III}	ξ_1^{IV}	ξ_1^{V}	ξ_1^{VI}	ξ_1^{VII}	ξ_1^{VIII}	ξ_1^{IX}	ξ_1^{X}	ξ_1^{XI}
π_1^{I}	π_1^{II}	π_1^{III}	π_1^{IV}	π_1^{V}	π_1^{VI}	π_1^{VII}	π_1^{VIII}	π_1^{IX}	π_1^{X}	π_1^{XI}
ρ_1^{I}	ρ_1^{II}	ρ_1^{III}	ρ_1^{IV}	ρ_1^{V}	ρ_1^{VI}	ρ_1^{VII}	ρ_1^{VIII}	ρ_1^{IX}	ρ_1^{X}	ρ_1^{XI}
σ_1^{I}	σ_1^{II}	σ_1^{III}	σ_1^{IV}	σ_1^{V}	σ_1^{VI}	σ_1^{VII}	σ_1^{VIII}	σ_1^{IX}	σ_1^{X}	σ_1^{XI}
τ_1^{I}	τ_1^{II}	τ_1^{III}	τ_1^{IV}	τ_1^{V}	τ_1^{VI}	τ_1^{VII}	τ_1^{VIII}	τ_1^{IX}	τ_1^{X}	τ_1^{XI}
υ_1^{I}	υ_1^{II}	υ_1^{III}	υ_1^{IV}	υ_1^{V}	υ_1^{VI}	υ_1^{VII}	υ_1^{VIII}	υ_1^{IX}	υ_1^{X}	υ_1^{XI}
ϕ_1^{I}	ϕ_1^{II}	ϕ_1^{III}	ϕ_1^{IV}	ϕ_1^{V}	ϕ_1^{VI}	ϕ_1^{VII}	ϕ_1^{VIII}	ϕ_1^{IX}		

[illegible][illegible]

CGTACAAACT	AATAAGAAAA	AGGTGACTGC	TCAAGAATTA	480
GATTACCTAA	CTCGTCACTA	TTTGGTGAAA	AATAAAAAAC	520
TCTATGAATT	TAACAACTCG	CCTTATGAAA	CGGGATATAT	560
TAAATTTATA	GAAAATGAGA	ATAGCTTTTG	GTATGACATG	600
ATGCCTGCAC	CAGGAGATAA	ATTTGACCAA	TCTAAATATT	640
TAATGATGTA	CAATGACAAT	AAAATGGTTG	ATTCTAAAGA	680
TGTGAAGATT	GAAGTTTATC	TTACGACAAA	GAAAAAGTGA	720
AATTATATTT	TAGAAAAGTA	AATATGAAGA	GTTAGTAATT	760
AAGGCAGGCA	CTTATAGAGT	ACCTGCCTTT	TCTAATATTA	800
TTTAGTTATA	GTTATTTTTG	TTATATCTCT	CTGATTTAGC	840
ATTAACCCCT	TGTTGCCATT	ATAGTTTTCAC	CAACTTTAG	880
CTGAAATTGG	GGGATCATTT	TTATCTTTAC	TATGGATAGT	920
TACTGTGTCG	CCGTTTTTAA	CGATTTGTTT	CTCTTTTAAT	960
TTGTCAGTTA	ATTTTTTCCA	TGCATCATTT	GCGTCAAACC	1000
TATTTCCATT	TGGATTTATT	CTTGACAAAT	CAATTCTTTT	1040
AACACTATCG	GTATTAATCG	GCTTGTATT	AAAATTACTA	1080
AGTTCATCTA	AATCAGCTGT	ACCCGTAATA	CTACTTTCGC	1120
CACCATTATT	TAAATTGTAC	GTAACACCAA	CTGTCTCATT	1160
TGCTGTTTTA	TCGATAATAT	TTGCTTCTTT	CAAAGCATCT	1200
CTTACATTTT	TCCATAAGTC	TCTATCTGTT	ATTCAGAAG	1240
CCTTTGCAAC	GTTATTAATA	CCATTATAAT	TTGAAGAAGA	1280
ATGAAAACCT	GAACCTACTG	TTGTTAAAAC	TAAAGCACTT	1320
GCTATCAATG	TTCTTGTTAA	TAGTTTTTTA	TTCATTTTAT	1360
TTTCTCCTAT	AACTTATTTG	CAATCGAT		1388

(11) INFORMATION FOR SEQUENCE ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 239

[illegible]

GTCCTTATTA TAGCCCTGCT TTTACAAAAG GGGAAAAAGT 320
 TGA CT TAAAC AAAAAAGAA CTAAAAAAG CCAACATACT 360
 AGCGAAGGAA CTTATATCCA TTTCCAAATA AGTGGCGTTA 400
 CAAATACTGA AAAATTACCT ACTCCAATAG AACTACCTTT 440
 AAAAGTTAAG GTTCATGGTA AAGATAGCCC CT TAAAGTAT 480
 GGGCCAAAGT TCGATAAAAA ACAATTAGCT ATATCAACTT 520
 TAGACTTTGA AATTCGTCAT CAGCTAACTC AAATACATGG 560
 ATTATATCGT TCAAGCGATA AAACGGGTGG TTATTGGAAA 600
 ATAACAATGA ATGACGGATC CACATATCAA AGTGATTTAT 640
 CTAAAAAGTT TGAATACAAT ACTGAAAAAC CACCTATAAA 680
 TATTGATGAA ATAAAAACTA TAGAAGCAGA AATTAATTAA 720
 TTTACCACTT T 731

(13) INFORMATION FOR SEQUENCE ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 234
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Unknown
- (D) TOPOLOGY: Unknown

(ii) Molecule type: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Asn Lys Lys Leu Leu Met Asn Phe Phe
5 10

Ile Val Ser Pro Leu Leu Leu Ala Thr Thr
15 20

Ala Thr Asp Phe Thr Pro Val Pro Leu Ser
25 30

Ser Asn Gln Ile Ile Lys Thr Ala Lys Ala
35 40

Ser Thr Asn Asp Asn Ile Lys Asp Leu Leu
45 50

Asp Trp Tyr Ser Ser Gly Ser Asp Thr Phe
55 60

1030242301

Thr	Asn	Ser	Glu	Val	Leu	Asp	Asn	Ser	Arg
				65					70
Gly	Ser	Met	Arg	Ile	Lys	Asn	Thr	Asp	Gly
				75					80
Ser	Ile	Ser	Leu	Ile	Ile	Phe	Pro	Ser	Pro
				85					90
Tyr	Tyr	Ser	Pro	Ala	Phe	Thr	Lys	Gly	Glu
				95					100
Lys	Val	Asp	Leu	Asn	Thr	Lys	Arg	Thr	Lys
				105					110
Lys	Ser	Gln	His	Thr	Ser	Glu	Gly	Thr	Tyr
				115					120
Ile	His	Phe	Gln	Ile	Ser	Gly	Val	Thr	Asn
				125					130
Thr	Glu	Lys	Leu	Pro	Thr	Pro	Ile	Glu	Leu
				135					140
Pro	Leu	Lys	Val	Lys	Val	His	Gly	Lys	Asp
				145					150
Ser	Pro	Leu	Lys	Tyr	Gly	Pro	Lys	Phe	Asp
				155					160
Lys	Lys	Gln	Leu	Ala	Ile	Ser	Thr	Leu	Asp
				165					170
Phe	Glu	Ile	Arg	His	Gln	Leu	Thr	Gln	Ile
				175					180
His	Gly	Leu	Tyr	Arg	Ser	Ser	Asp	Lys	Thr
				185					190
Gly	Gly	Tyr	Trp	Lys	Ile	Thr	Met	Asn	Asp
				195					200
Gly	Ser	Thr	Tyr	Gln	Ser	Asp	Leu	Ser	Lys
				205					210
Lys	Phe	Glu	Tyr	Asn	Thr	Glu	Lys	Pro	Pro
				215					220
Ile	Asn	Ile	Asp	Glu	Ile	Lys	Thr	Ile	Glu
				225					230
Ala	Glu	Ile	Asn						

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1095
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Unknown
(D) TOPOLOGY: Unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

ATCATTAAAT	ATAATTAATT	TTCTTTTAAAT	ATTTTTTTTAA	40
TTGAATATTT	AAGATTATAA	GATATATTTA	AAGTGTATCT	80
AGATACTTTT	TGGGAATGTT	GGATGAAGGA	GATAAAAATG	120
AATAAGAGTC	GATTTATTTT	ATGCGTAATT	TTGATATTCG	160
CACTTATACT	AGTTCTTTTT	ACACCCAACG	TATTAGCAGA	200
GAGCCAACCA	GACCCTACGC	CAGATGAGTT	GCACAAAGCG	240
AGTAAATTCA	CTGGTTTGAT	GGAAAATATG	AAAGTTTTAT	280
ATGATGATCA	TTATGTATCA	GCAACTAAAG	TTAAGTCTGT	320
AGATAAATTT	AGGGCACATG	ATTTAATTTA	TAACATTAGT	360
GATAAAAAAC	TGAAAAATTA	TGACAAAGTG	AAAAACAGAGT	400
TATTAAATGA	AGGTTTAGCA	AAGAAGTACA	AAGATGAAGT	440
AGTTGATGTG	TATGGATCAA	ATTACTATGT	AAACTGCTAT	480
TTTTTCATCCA	AAGATAATGT	AGGTAAAGTT	ACAGGTGGCA	520
AAACTTGTAT	GTATGGAGGA	ATAACAAAAC	ATGAAGGAAA	560
CCACTTTGAT	AATGGGAACT	TACAAAATGT	ACTTATAAGA	600
GTTTATGAAA	ATAAAAGAAA	CACAATTTCT	TTTGAAGTGC	640
AAACTGATAA	GAAAAGTGTA	ACAGCTCAAG	AACTAGACAT	680
AAAAGCTAGG	AATTTTTTTAA	TTAATAAAAA	AAATTTGTAT	720
GAGTTTAAAC	GTTCAACCATA	TGAAACAGGA	TATATAAAAT	760
TTATTGAAAA	TAACGGCAAT	ACTTTTTTGGT	ATGATATGAT	800

(15) INFORMATION FOR SEQUENCE ID NO:14:

(A) LENGTH: 266
(B) TYPE: Amino Acid
(C) STRANDEDNESS: Unknown
(D) TOPOLOGY: Unknown

(ii) Molecule type: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met	Asn	Lys	Ser	Arg 5	Phe	Ile	Ser	Cys	Val 10
Ile	Leu	Ile	Phe	Ala 15	Leu	Ile	Leu	Val	Leu 20
Phe	Thr	Pro	Asn	Val 25	Leu	Ala	Glu	Ser	Gln 30
Pro	Asp	Pro	Thr	Pro 35	Asp	Glu	Leu	His	Lys 40
Ala	Ser	Lys	Phe	Thr 45	Gly	Leu	Met	Glu	Asn 50
Met	Lys	Val	Leu	Tyr 55	Asp	Asp	His	Tyr	Val 60
Ser	Ala	Thr	Lys	Val 65	Lys	Ser	Val	Asp	Lys 70
Phe	Arg	Ala	His	Asp 75	Leu	Ile	Tyr	Asn	Ile 80

Ser	Asp	Lys	Lys	Leu 85	Lys	Asn	Tyr	Asp	Lys 90
Val	Lys	Thr	Glu	Leu 95	Leu	Asn	Glu	Gly	Leu 100
Ala	Lys	Lys	Tyr	Lys 105	Asp	Glu	Val	Val	Asp 110
Val	Tyr	Gly	Ser	Asn 115	Tyr	Tyr	Val	Asn	Cys 120
Tyr	Phe	Ser	Ser	Lys 125	Asp	Asn	Val	Gly	Lys 130
Val	Thr	Gly	Gly	Lys 135	Thr	Cys	Met	Tyr	Gly 140
Gly	Ile	Thr	Lys	His 145	Glu	Gly	Asn	His	Phe 150
Asp	Asn	Gly	Asn	Leu 155	Gln	Asn	Val	Leu	Ile 160
Arg	Val	Tyr	Glu	Asn 165	Lys	Arg	Asn	Thr	Ile 170
Ser	Phe	Glu	Val	Gln 175	Thr	Asp	Lys	Lys	Ser 180
Val	Thr	Ala	Gln	Glu 185	Leu	Asp	Ile	Lys	Ala 190
Arg	Asn	Phe	Leu	Ile 195	Asn	Lys	Lys	Asn	Leu 200
Tyr	Glu	Phe	Asn	Ser 205	Ser	Phe	Tyr	Glu	Thr 210
Gly	Tyr	Ile	Lys	Phe 215	Ile	Glu	Asn	Asn	Gly 220
Asn	Thr	Phe	Trp	Tyr 225	Asp	Met	Met	Pro	Ala 230
Pro	Gly	Asp	Lys	Phe 235	Asp	Gln	Ser	Lys	Tyr 240
Leu	Met	Met	Tyr	Asn 245	Asp	Asn	Lys	Thr	Val 250
Asp	Ser	Lys	Ser	Val 255	Lys	Ile	Glu	Val	His 260

Leu Thr Thr Lys Asn Gly
265

(16) INFORMATION FOR SEQUENCE ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1837
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Unknown
(D) TOPOLOGY: Unknown

(ii) Molecule type: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TCATGTTTGA	CAGCTTATCA	TCGATAAGCT	TACTTTTTCGA	40
ATCAGGTCTA	TCCTTGAAAC	AGGTGCAACA	TAGATTAGGG	80
CATGGAGATT	TACCAGACAA	CTATGAACGT	ATATACTCAC	120
ATCACGCAAT	CGGCAATTGA	TGACATTGGA	ACTAAATTCA	160
ATCAATTTGT	TACTAACAAG	CAACTAGATT	GACAACTAAT	200
TCTCAACAAA	CGTTAATTTA	ACAACATTCA	AGTAACTCCC	240
ACCAGCTCCA	TCAATGCTTA	CCGTAAGTAA	TCATAACTTA	280
CTAAAACCTT	GTTACATCAA	GGTTTTTTTCT	TTTTGTCTTG	320
TTCATGAGTT	ACCATAACTT	TCTATATTAT	TGACAACTAA	360
ATTGACAACT	CTTCAATTAT	TTTTCTGTCT	ACTCAAAGTT	400
TTCTTCATTT	GATATAGTCT	AATTCCACCA	TCACTTCTTC	440
CACTCTCTCT	ACCGTCACAA	CTTCATCATC	TCTCACTTTT	480
TCGTGTGGTA	ACACATAATC	AAATATCTTT	CCGTTTTTAC	520
GCACTATCGC	TACTGTGTCA	CCTAAAATAT	ACCCCTTATC	560
AATCGCTTCT	TTAAACTCAT	CTATATATAA	CATATTTTCAT	600
CCTCCTACCT	ATCTATTTCGT	AAAAAGATAA	AAATAACTAT	640
TGTTTTTTTTT	GTTATTTTAT	AATAAAATTA	TTAATATAAG	680
TTAATGTTTTT	TTAAAAATAT	ACAATTTTAT	TCTATTTATA	720

GTTAGCTATT	TTTTCATTGT	TAGTAATATT	GGTGAATTGT	760
AATAACCTTT	TTAAATCTAG	AGGAGAACCC	AGATATAAAA	800
TGGAGGAATA	TTAATGGAAA	ACAATAAAAA	AGTATTGAAG	840
AAAATGGTAT	TTTTTGTTTT	AGTGACATTT	CTTGGACTAA	880
CAATCTCGCA	AGAGGTATTT	GCTCAACAAG	ACCCCGATCC	920
AAGCCAACTT	CACAGATCTA	GTTTAGTTAA	AAACCTTCAA	960
AATATATATT	TTCTTTATGA	GGGTGACCCT	GTTACTCACG	1000
AGAATGTGAA	ATCTGTTGAT	CAACTTAGAT	CTCACGATTT	1040
AATATATAAT	GTTTCAGGGC	CAAATTATGA	TAAATTAAAA	1080
ACTGAACTTA	AGAACCAAGA	GATGGCAACT	TTATTTAAGG	1120
ATAAAAACGT	TGATATTTAT	GGTGTAGAAT	ATTACCATCT	1160
CTGTTATTTA	TGTGAAAATG	CAGAAAGGAG	TGCATGTATC	1200
TACGGAGGGG	TAACAAATCA	TGAAGGGAAT	CATTTAGAAA	1240
TTCCATAAAA	GATAGTCGTT	AAAGTATCAA	TCGATGGTAT	1280
CCAAAGCCTA	TCATTTGATA	TTGAAACAAA	TAAAAAAATG	1320
GTAAC TGCTC	AAGAATTAGA	CTATAAAGTT	AGAAAATATC	1360
TTACAGATAA	TAAGCAACTA	TATACTAATG	GACCTTCTAA	1400
ATATGAAACT	GGATATATAA	AGTTCATACC	TAAGAATAAA	1440
GAAAGTTTTT	GGTTTGATTT	TTTCCCTGAA	CCAGAATTTA	1480
CTCAATCTAA	ATATCTTATG	ATATATAAAG	ATAATGAAAC	1520
GCTTGACTCA	AACACAAGCC	AAATTGAAGT	CTACCTAACA	1560
ACCAAGTAAC	TTTTTGCTTT	TGGCAACCTT	ACCTACTGCT	1600
GGATTTAGAA	ATTTTATTGC	AATTCTTTTA	TTAATGTAAA	1640
AACCGCTCAT	TTGATGAGCG	GTTTTGTCTT	ATCTAAAGGA	1680
GCTTTACCTC	CTAATGCTGC	AAAATTTTAA	ATGTTGGATT	1720
TTTGTATTTG	TCTATTGTAT	TTGATGGGTA	ATCCCATTTT	1760
TCGACAGACA	TCGTCGTGCC	ACCTCTAACA	CCAAAATCAT	1800

1837

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 251
(B) TYPE: Amino Acid
(C) STRANDEDNESS: Unknown
(D) TOPOLOGY: Unknown

(ii) Molecule type: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met Glu Asn Asn Lys Lys Val Leu Lys Lys
5 10

Met Val Phe Phe Val Leu Val Thr Phe Leu
15 20

Gly Leu Thr Ile Ser Gln Glu Val Phe Ala
25 30

Gln Gln Asp Pro Asp Pro Ser Gln Leu His
35 40

Arg Ser Ser Leu Val Lys Asn Leu Gln Asn
45 50

Ile Tyr Phe Leu Tyr Glu Gly Asp Pro Val
55 60

Thr His Glu Asn Val Lys Ser Val Asp Gln
65 70

Leu Arg Ser His Asp Leu Ile Tyr Asn Val
75 80

Ser Gly Pro Asn Tyr Asp Lys Leu Lys Thr
85 90

Glu Leu Lys Asn Gln Glu Met Ala Thr Leu
95 100

Phe Lys Asp Lys Asn Val Asp Ile Tyr Gly
105 110

Val Glu Tyr Tyr His Leu Cys Tyr Leu Cys
115 120

Glu Asn Ala Glu Arg Ser Ala Cys Ile Tyr
125 130

Gly Gly Val Thr Asn His Glu Gly Asn His

•

SpeA reverse primer, adds SpeB overlap

5' GAGATTTAACAACCTGGTTGCTTGGTTGTTAGGTAGAC 3' 37-mer

3. SpeB forward primer, adds SpeA overlap:

5' GTCTACCTAACAACCAAGC A A C C A G T T G T T A A A T C T C 3' 37-mer

4. SpeB reverse primer; adds stop site (Amber) and maintains BamHI site:

5' GAATTCGGATCCGCTAGCCTACAACAG 3' 27-mer

SpeA (L42R) -SpeB (C47S) gene insert DNA sequence

1936	1937	1938	1939	1940	1941	1942	1943	1944	1945	1946	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344</
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	--------

SEQ ID NO:24

1	MNKKKL	GIRL	L	SLLALGGFV	LANPVFADQN	FARNEKEAKD	SAITFIQKSA	AIKAGARSAE
61	DIKLDKVN	LG	GELSGSNMYV	YNISTGGFVI	VSGDKRSPEI	LGYSTSGSFD	ANGKENIASF	
121	MESYVEQ	IK	NKKLDTTYAG	TAEIKQPVVK	SLLDSKGIHY	NQGNPYNLLT	PVIEKVKPGE	
181	QSFVGQHA	AT	GCVATATAQI	MKYHNYPNKG	LKDYTYTLSS	NNPYFNHPKN	LFAAISTRQY	
241	NWNNILPT	YS	GRESNVQKMA	ISELMAQVGI	SVDMDYGPSS	GSAGSSRVQR	ALKENFGYNQ	
301	SVHQINRG	DF	SKQDWEAQID	KELSQNPVY	YQGVKGKVGH	AFVIDGADGR	NFYHVNWGWG	
361	GVSDGFFR	LD	ALNPSALGTG	GGAGGFNGYQ	SAVVGIKP			

SpeB clone used for fusion, mature polypeptide. Estimated M_r = 28.75 kDa

1	MQQDPDPSQL	HRSSLVKNLQ	NIYFLYEGDP	VTHENVKSVD	QLRSHDLIYN
51	VSGPNYDKLK	TELKNQEMAT	LFKDKNIDIY	GVEYYHLCYL	CENAERSACI
101	GGVTNREGNH	LEIPKKIVVK	VSIDGIQSLS	FDIETNKKMV	TAQELDYKVR
151	KYLTDNKQLY	TNGPSKYETG	YIKFIPKNKE	SFWFDFPPEP	EFTQSKYLM
201	YKDNETLDSN	TQIEVYLTTK	QPVVKSL LDS	KGIHYNQGNP	YNLLTPVIEK
251	VKPGEQSFVG	QHAATGCVAT	ATAQIMKYHN	YPNKGLKDYT	YTLSSNNPYF
301	NHPKNLFAAI	STRQYNWNNI	LPTYSGRESN	VQKMAISELM	ADVGISVDM
351	YGPSSGSAGS	SRVQRALKEN	FGYNQSVHQI	NRSDFSQDWE	AQIDKELSQN
401	QPVYYQGGKV	GGHAFVIDGA	DGRNFIYHVN	GWGGVSDGFF	RLDALNPSAL
551	GTGGGAGGFN	GYOSAVVG			

SpeA [L42R]-SpeB [C47S] fusion. Estimated M_r = 54 kDa

1	MQQDPDPSQL	HRSSLVKNLQ	NIYFLYEGDP	VTHENVKSVD	QLRSHDLIYN
51	VSGPNYDKLK	TELKNQEMAT	LFKDKNIDIY	GVEYYHLCYL	CENAERSACI
101	GGVTNREGNH	LEIPKKIVVK	VSIDGIQSLS	FDIETNKKMV	TAQELDYKVR
151	KYLTDNKQLY	TNGPSKYETG	YIKFIPKNKE	SFWFDFPPEP	EFTQSKYLM
201	YKDNETLDSN	TQIEVYLTTK	QPVVKSL LDS	KGIHYNQGNP	YNLLTPVIEK
251	VKPGEQSFVG	QHAATGCVAT	ATAQIMKYHN	YPNKGLKDYT	YTLSSNNPYF
301	NHPKNLFAAI	STRQYNWNNI	LPTYSGRESN	VQKMAISELM	ADVGISVDM
351	YGPSSGSAGS	SRVQRALKEN	FGYNQSVHQI	NRSDFSQDWE	AQIDKELSQN
401	QPVYYQGGKV	GGHAFVIDGA	DGRNFIYHVN	GWGGVSDGFF	RLDALNPSAL
551	GTGGGAGGFN	GYOSAVVG			

1. SpeA forward primer, including NdeI site:

SEO ID NO:29

2. **SpeA reverse primer**; kills NdeI site, adds SpeB overlap:

5' CATGTGTATATCTCCTTCCTTGGTTGTTAGGTAGAC 3' 36-mer

SEQ ID NO:30

3. SpeB forward primer; kills NdeI site, adds SpeA overlap:

5' GTCTACCTAACAACCAAGGAAGGAGATATACACATG 3' 36-mer

SEO ID NO:31

4. SpeB reverse primer; adds stop site (Amber) and maintains BamHI site:

5' GAATTCGGATCCGCTAGCCTACAACAG 3' 27-mer